AMERICAN HOUSING SURVEY

Sample Case History File: 2015 and Beyond

PD&R

Contents

1. Overview

The purpose of this document is to describe the American Housing Survey (AHS) *National Sample Case History File*. This file tracks all integrated national sample and independent metropolitan area sample AHS cases that were interviewed beginning in 2015[1] and provide basic information on the interview status.

For 2015, the U.S. Department of Housing and Urban Development (HUD) and the U.S. Census Bureau drew an entirely new sample for the AHS. The 2015 AHS longitudinal sample is composed of an integrated national longitudinal sample and 20 independent metropolitan area longitudinal oversamples. The national longitudinal sample is described as integrated because it incorporates a few different types of samples. The integrated national longitudinal sample includes:

- A representative sample of the nation.
- Representative oversamples of each of the top 15 group of metropolitan areas[2].
- A representative oversample of HUD-assisted housing units.

The entire integrated national longitudinal sample will be surveyed every two years, meaning it is a two-year survey cycle.

In 2015, HUD and the Census Bureau identified a "Next 20" group of metropolitan areas to be surveyed in 2015 and later years. These Next 20 group are described as independent so as not to be confused with the Top 15 metropolitan areas that are part of the integrated national longitudinal sample. The Next 20 group is a set of metropolitan areas ranging from the 16th to 51st largest by population, and it includes: [3]

The first ten of the Next 20 group were surveyed in 2015 and 2019. They are scheduled to be surveyed every four years thereafter (2023, etc.). The second ten of the Next 20 group of metropolitan area longitudinal oversamples were surveyed in 2017 and are scheduled to be surveyed every four years thereafter (2021, 2025, etc.).

During data collection, cases are categorized as completed interviews or noninterviews. Noninterviews are classified into three groups: type A, B, and C. Type A noninterviews are sample addresses that meet the definition of a housing unit where an interview is unable to be conducted. Type B noninterviews are sample addresses that do not meet the definition of a housing unit, but could become eligible in the future. Type C non-interview classifications occur when structural changes have occurred that result in a change of address, or due to incomplete or bad addresses. Type C cases will not be interviewed in

---

[1] The *National Sample Case History File* for survey years 1985 to 2013 can be found at https://www.census.gov/programs-surveys/ahs/tech-documentation/help-guides/1985-2013/national-sample-case-history_2013.html

[2] The Top 15 group of metropolitan area longitudinal oversamples use the 2013 Office of Management and Budget's core based statistical area definitions as of February 2013

[3] For more information about how the Next 20 group of metropolitan areas was selected, see *Metropolitan Area Selection Strategy: 2015 and Beyond* at https://www.census.gov/programs-surveys/ahs/operations-and-administration/redesign-2015/2015-metro-area-selection.html

subsequent AHS data collection periods. This document will refer to completed interviews, type As, and type Bs collectively as "eligible housing units" and type Cs as "ineligible housing units".

As of 2019, the *Sample Case History File* has 150,910 cases. This is the total number of unique cases that were sampled beginning in 2015. This is a flat file structure, meaning there is one record for each sample case that was interviewed, which can be uniquely identified by the *CONTROL* variable. This file contains a variable for each year with the Weight *PWT*, noninterview code *NOINT*, interview status *INTSTATUS*, *TENURE*, year the case became a noninterview *YEARMC*, 2013 Office of Management and Budget (OMB) definition of Core Based Statistical Area (CBSA) code *OMB13CBSA*, and year the case was introduced to the sample *YRINTRO* variables.

2. Data Dictionary

CONTROL

> Control number

Description

> This is the unique identifier for each record on the file and can be used to merge this file with the AHS public use file (PUF).

Type:

> Character

PWTXX

> Pure weight, or the inverse of the probability of selection

Values

> Missing–Case was not on the file in this year

> N–Case was not in sample this year, returning type C

> 1:9999.9999

Description

> *PWT* is the product of base weight, weighting control factors, and sample adjustment factor. The last two digits of the variable name correspond to the last two digits of the year to which it refers. For example, *PWT15* represents the *PWT* value from 2015.

Type

> Numeric

NOINTXX

> Reason for noninterview

Values

> Missing—Case was not on the file in this year

> N—Not applicable, interviewed case

1—Type A: No one home

2—Type A: Temporarily absent

3—Type A: Refusal

4—Type A: Unable to locate

5—Type A: Language problem

6—Type A: Other, occupied

10—Type B: Permit granted, construction not started

11—Type B: Under construction, not ready

12—Type B: Permanent or temporary business or commercial storage

13—Type B: Unoccupied site for mobile home or tent

14—Type B: Other unit including nonstaff, or converted to institutional unit

15—Type B: Occupancy prohibited

16—Type B: Interior exposed to the elements

17—Type B: Not classified above, structure type is not boat, RV, tent, cave, or railroad car

18—Type B: Not classified above, structure type is boat, RV, tent, cave, or railroad car

19—Type B: Demolished or disaster loss (2019 and beyond)

30—Type C: Demolished or disaster loss (2015 and 2017)

31—Type C: House or mobile home moved

32—Type C: Unit eliminated in structural conversion

33—Type C: Merged, not in current sample

36—Type C: Permit abandoned

37—Type C: Not classified above

42—Type C: Unit does not exist, or unit is out of scope

43—Type C: Unable to locate sample address

Description

This code provides the reason a case was a noninterview. The last two digits of the variable name correspond to the last two digits of the year to which it refers. For example, *NOINT15* represents the *NOINT* value from 2015.

Note: Type C "demolished or disaster loss" noninterviews (NOINT = 30) were reclassified to Type B noninterviews (NOINT = 19) in the 2019 AHS. To be consistent with the reclassification, 261 national sample cases from the 2017 AHS were identified as demolished or disaster losses, and became eligible for interview again in 2019.

Type

Numeric

OMB13CBSA

2013 OMB CBSA code

Values

12060—Atlanta-Sandy Springs-Roswell, GA

12580—Baltimore-Columbia-Towson, MD

13820—Birmingham-Hoover, AL

14460—Boston-Cambridge-Newton, MA-NH

16980—Chicago-Naperville-Elgin, IL-IN-WI

17140—Cincinnati, OH-KY-IN

17460—Cleveland-Elyria, OH

19100—Dallas-Fort Worth-Arlington, TX

19740—Denver-Aurora-Lakewood, CO

19820—Detroit-Warren-Dearborn, MI

26420—Houston-The Woodlands-Sugar Land, TX

28140—Kansas City, MO-KS

29820—Las Vegas-Henderson-Paradise, NV

31080—Los Angeles-Long Beach-Anaheim, CA

32820—Memphis, TN-MS-AR

33100—Miami-Fort Lauderdale-West Palm Beach, FL

33340—Milwaukee-Waukesha-West Allis, WI

33460—Minneapolis-St. Paul-Bloomington, MN-WI

35380—New Orleans-Metairie, LA

35620—New York-Newark-Jersey City, NY-NJ-PA

36420—Oklahoma City, OK

37980—Philadelphia-Camden-Wilmington, PA-NJ-DE-MD

38060—Phoenix-Mesa-Scottsdale, AZ

38300—Pittsburgh, PA

38900—Portland-Vancouver-Hillsboro, OR-WA

39580—Raleigh, NC

40060—Richmond, VA

40140—Riverside-San Bernardino-Ontario, CA

40380—Rochester, NY

41700—San Antonio-New Braunfels, TX

41860—San Francisco-Oakland-Hayward, CA

41940—San Jose-Sunnyvale-Santa Clara, CA

42660—Seattle-Tacoma-Bellevue, WA

45300—Tampa-St. Petersburg-Clearwater, FL

47900—Washington-Arlington-Alexandria, DC-VA-MD-WV

99998—All other metropolitan areas

99999—Not in a metropolitan area

Description

These values adhere to the OMB definitions for core based statistical areas (CBSAs). More information on these can be found at https://www.census.gov/programs-surveys/metro-micro.html. *OMB13CBSA* codes are only produced for cases in the Top 15 group of metropolitan area longitudinal oversamples[4] (on the national microdata) and cases on the metropolitan microdata files.

Type

Character

INTSTATUSXX

Interview status

Values

Missing—Case was not on the file in this year

1—Occupied interview

2—URE (Usual Residence Elsewhere) interview

3—Vacant interview

4—Noninterview

---

[4] The Top 15 group of metropolitan area longitudinal oversamples use the 2013 Office of Management and Budget's core based statistical area definitions as of February 2013

Description

This code provides the interview status of a case for a given year. The last two digits of the variable name correspond to the last two digits of the year to which it refers. For example, *INTSTATUS15* represents the *INTSTATUS* value shown on the 2015 PUF.

Type

Character

TENUREXX

Owner or renter status of unit

Values

Missing—Case was not on the file in this year

N—Not applicable

1—Owned or being bought by someone in your household

2—Rented

3—Occupied without a payment of rent

Description

This code provides the owner or renter status of the case for a given year. The last two digits of the variable name correspond to the last two digits of the year to which it refers. For example, *TENURE15* represents the *TENURE* value shown on the 2015 PUF.

Type

Character

YEARMC

Year made Type C

Values

Missing—Case was not made a type C

2015:20XX—2015–20XX

Description

This refers to the year the case was made an ineligible housing unit (a type C). If the case becomes an eligible housing unit and later becomes ineligible again, the first year it was shown on the PUF as an ineligible housing unit is shown.

Type

Character

YRINTRO

Year introduced

Values

2015—20XX –2015–20XX

Description

This refers to the year the case was first introduced to the sample. If the case moved in and out of the sample, the first year the case is on a PUF is shown.

Type

Character

SAMPLETYPE

(Available 2017 and later)

Sample designation

Values

1—Integrated national longitudinal sample

2—First wave of Next 20 group of metropolitan area longitudinal oversamples[5]

3—Second wave of Next 20 group of metropolitan area longitudinal oversamples

Description

This indicates the sample designation to which a case belongs. The integrated national longitudinal sample includes cases that are interviewed by design in every interview year. This includes national longitudinal sample cases and Top 15 (static) metropolitan sample cases. The first wave of metropolitan areas includes 10 metropolitan areas that were interviewed in 2015, 2019, and will be interviewed every four years thereafter. The second wave of metropolitan areas includes 10 metropolitan areas that were interviewed in 2017, are schedule to be interviewed in 2021, and will be interviewed every four years thereafter.

Type

Character

---

[5] The Next 20 group of metropolitan area longitudinal oversamples use the 2013 Office of Management and Budget's core based statistical area definitions as of February 2013.

U.S. Department of Housing and Urban Development

Office of Policy Development and Research

Washington, DC 20410-6000